Remote Ambulatory Management of Veterans with Sleep Apnea

Identifier: NCT03007745

Document data: 10/20/2017

RESEARCH DESIGN AND METHODS

Aim 1 Effect of REVAMP management compared to in-person care on functional outcomes

Study population and enrollment of participants. Veterans (n=356) referred to one of the sleep centers for evaluation of OSA will be recruited. Men and women will be included with no restrictions based on race and ethnicity. We will target women's clinics to enhance the recruitment of women. Healthcare providers submitting referrals will be asked on the CPRS consult form whether or not the patient is interested in participating in the study. Interested patients will be contacted by phone *by the study coordinator to explain the study and determine if they have internet access and are eligible based on the inclusion/exclusion criteria. Eligible patients interested in participation will be scheduled to meet with the study coordinator to obtain signed informed consent.* Patients who are ineligible or decline to participate will be scheduled for a routine initial clinic appointment. Enrolled participants will be given the uniform resource locator (url) address to the REVAMP website and a user name and temporary password. The participant's primary care provider will be informed that his/her patient is enrolled in the study and that all clinically indicated care should continue to be provided.

Eligibility criteria and enrollment. Veterans must meet the following inclusion criteria prior to enrollment:

1. Age ≥ 18 years and enrolled at one of the participating VAMCs
2. Referral to one of the participating sleep centers for evaluation of suspected OSA
3. Access in the home to the internet, e-mail, and phone on all days
4. *Living within a 45 miles radius of the clinical site so that participants randomized to in-person delivery of care will be able to travel to and from the sleep center*
5. Fluent in English as assessed on the initial phone contact

Individuals will be excluded from the study for the following reasons:

1. Unable or unwilling to provide informed consent and complete required questionnaires
2. Previous diagnosis of OSA, central sleep apnea (≥ 50% of apneas on diagnostic testing are central apneas), Cheyne-Stokes breathing, obesity hypoventilation syndrome, or narcolepsy
3. Previous treatment with PAP, non-nasal surgery for OSA, or current use of supplemental oxygen
4. A clinically unstable medical condition in the previous 2 months as defined by a new medical diagnosis (e.g., pneumonia, cardiac disease, thyroid disease, depression or psychosis, cirrhosis, cancer, etc.)
5. Individuals who regularly experience jet lag or irregular work schedules over the previous 3 months
6. Women who are pregnant or women who are sexually active and of child-bearing age who are not using some form of contraceptive
7. *Unable to perform tests due to visual (inability to read the consent form) or hearing (unable to have a phone conversation) impairment, or unable to perform activities of daily living*

Participants will complete the following baseline questionnaires on the website: FOSQ-10,[47] EES,[48, 49] SF-12,[50, 51] CES-D,[52] EQ-5D (including the EuroQol feeling thermometer),[7, 8] and medical service use not captured by their electronic medical record (Table 2 and Appendix 3). The Charlson comorbidity index[55] will be used to determine if health status predicts which patients are good candidates for REVAMP management.

To monitor for possible selection bias, all potential participants will be tracked to determine the number who agree to be contacted, the number who provide informed consent, the criteria that excluded an individual's participation, and the reasons why participants were unable to complete the protocol. Post-hoc analysis will compare patient characteristics at baseline to assess for any differences across groups and sites. Following completion of the baseline questionnaires, the CES-D score will be reviewed by a member of the research team who is not involved with delivery of care. The participant's primary care provider will be notified if the CES-D score is greater than 24, indicating the presence of severe depression. These participants will be allowed to continue in the study since all participants with OSA will receive APAP treatment and this intervention has been shown to improve depression.[5, 56] To assess for any medication changes that might affect functional outcomes, a list of the participant's medications, with their dosage and frequency, will be obtained from the patient and the patient's medical record during the first 3 months of PAP treatment.

All participants will complete the website questionnaires in order to standardize collection of our functional outcomes. Practitioners in both arms will not have access to the responses to the patient satisfaction questionnaires (WAI-SR and CSQ-8). The other questionnaire responses of the participants randomized to REVAMP management will be available to the practitioners performing the phone clinic evaluations. Enrolled subjects assigned to the in-person arm will only have access to the website's questionnaires but none of its other features. They will be informed that a research staff member will contact them to schedule their first clinic appointment. None of the questionnaires they complete on the website will be shared with their practitioners. At the time of their in-person clinic visits, these participants will complete the questionnaires that are routinely used as a part of the clinical site's usual care. That information will not be collected for analysis.

Table 2. Data collection and action schedule for all participants						
Measures & Procedures	Baseline assessment	Initial clinical assessment	AutoCPAP set-up	1-week phone FU	1-month FU assessment	3-month FU assessment
	Week 0	Week 2-4	Week 8	Week 9-12	Week 12	Week 20
Eligibility criteria	In-person					
Informed consent	In-person					
Patient characteristic questionnaire	website					
Functional & QOL questionnaires†	website				website	Website $^{\Delta}$
Therapy process indicators‡		website		website	website	Website $^{\Delta}$
Clinical assessment*		In-person vs phone		Phone	In-person vs phone	In-person vs phone
APAP set-up			DME			
APAP adherence*				Wireless	Wireless	Wireless

DME = home health care company; FU=follow-up; QO =quality of life; Δ=performed every 3 months thereafter to the end of data collection

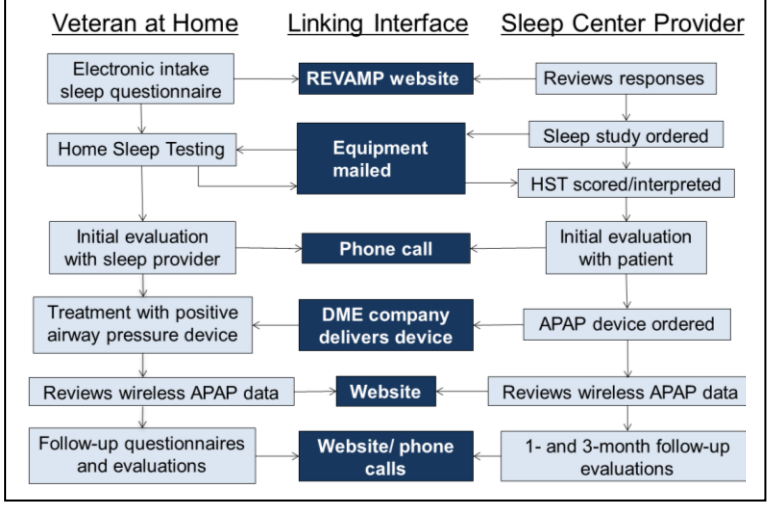
* Subjects randomized to in-person management will have in-person clinic evaluations in the sleep center and subjects randomized to REVAMP management will have phone clinic sessions. APAP adherence in both groups will be obtained by wireless modem.

† **Functional & QOL questionnaires: FOSQ-10, ESS, SF-12, CES-D, and EQ-5D**; ‡ Therapy process indicators: **WAI, CSQ-8** (performed within 1-week following each encounter)

Participant randomization. Upon completing baseline questionnaires, a computer-generated, blocked randomization scheme in the REVAMP program will automatically allocate eligible participants into one of two groups: *REVAMP management versus in-person management.* Randomization at each site will be programmed by SAS module Proc Plan. Block size at each site will be six subjects so that among each block of six randomized individuals, exactly three will be assigned to each group in a random order. Investigation and management of the two arms of care will be conducted over equal time periods.

REVAMP pathway. Following completion of the website questionnaires and the HST, REVAMP participants will be scheduled for an initial phone interview with the *sleep specialist (MD or NP)* (Fig 4). The practitioner will review the results with the subject. *Using the criteria required by the Center of Medicare and Medicaid Services, participants with AHI ≥ 15 events/hr and subjects with 5 ≤AHI <15 who have OSA-related symptoms will be diagnosed with OSA.* Those interested in PAP treatment will be prescribed an APAP unit with wireless modem. Subjects with OSA who are found on follow-up to require bilevel PAP and/or supplemental oxygen will continue to participate in Aim 1. Participants testing negative for OSA will be offered usual clinical care to further evaluate their sleep-related symptoms. Participants with a diagnosis of central sleep apnea/Cheyne-Stokes respiration (CSA/CSR) will be referred to a sleep physician and PSG testing. *Participants without OSA on HST and participants with CSA/ CSR will not be included in the Aim 1 modified intent to treat analysis set but will be included in the primary analysis set of Aim 2 and Aim 3.*

Fig 4. Interfaces linking Veteran and sleep center provider in REVAMP



The wireless modem on the home PAP unit sends PAP data daily to the manufacturer's server which in turn exports the data daily to the REVAMP website. The website displays the following information in an easily understandable format: hours of use, amount of air leak, and AHI (Fig. 2). Participants and practitioners can view the data on the REVAMP website. Participants will be instructed to review their results on the website daily. The number of times a subject logs into the website will be tracked.

Following the first week of PAP treatment, a therapist with at least one year of experience managing patients with OSA will phone the participant. Using a standardized protocol, the therapist will evaluate the participant's response to treatment, inquire if any barriers are limiting PAP use, and provide advice on how to overcome those problems (*see Appendix 4*). The therapist will encourage the participant to review his/her PAP data on the website. The participant will be scheduled for 1- and 3-month follow-up phone calls with the therapist. Additional phone calls are allowed as clinically indicated and will be tracked. The therapist will have the option of consulting with the sleep specialist (MD or NP). Upon consultation, if it is felt that the problem cannot be addressed by phone, the participant will be scheduled for an in-person clinic appointment with the physician. The type, number

and duration of all encounters will be recorded in case report forms. All phone clinic sessions and other patient encounters will be documented in the participant's electronic medical record.

REVAMP has a secure message system that sends e-mail and text messages to the participant alerting him/her to login to the website to view a message. The REVAMP website will use this feature to remind the participant to complete questionnaires, view their PAP data, and keep scheduled phone clinic appointments.

In-person management. *Participants in the in-person arm will be scheduled to meet with the sleep therapist for in-person set-up instructions for HST. Following HST, a sleep specialist (MD or NP) will conduct the initial in-person clinic evaluation. Participants diagnosed with OSA (using the same criteria detailed above for the REVAMP pathway) who are willing to start PAP treatment will be prescribed APAP. Sleep therapists with at least one year of experience managing patients with OSA will perform the follow-up clinic evaluations. Participants will receive a 1-week phone call and be scheduled for 1- and 3-month in-person follow-up clinic visits. Additional phone calls and clinic visits will be allowed as clinically indicated. The type, number and duration of all encounters will be monitored and documented in the participant's electronic medical record*.

Other than the above requirements, each site will conduct in-person management according to their usual clinical practice. The decision to allow each site to conduct its current in-person clinical care will make our findings more generalizable. Not having to replicate the exact same conditions for in-person management at the sites reduces the cost of the project. We realize that allowing these differences at the 3 sites raises the possibility of a site-related bias. We will assess this by including site as a covariate in the analysis models.

Home sleep testing and polysomnogram. Participants in both the in-person and REVAMP arms will have an HST as their first diagnostic test, but their practitioners will have the ability to order an in-lab PSG when clinically indicated. A technically acceptable HST will require at least a 3 hr recording of oxygen saturation and one respiratory signal. Recordings that are technically inadequate will be repeated once. If the repeat HST is not of acceptable quality, the participant will be scheduled for an in-lab PSG *performed according to AASM recommended guidelines.*[57] To reduce the possibility of a false negative HST result, participants with an AHI < 5 events/hr on initial HST will repeat the study. Participants with two negative HSTs will be advised that they may still have OSA and, if clinically indicated, can be offered the option of an in-laboratory PSG.

The clinical sites will use a type 3 portable monitor that records nasal pressure, rib cage and abdominal movement, pulse oximetry, and body position. *Two of the sites use the T3 monitor (Nox Medical) and the third site uses Nomad, Nihon Kohden.* We will control for diagnostic bias by ensuring that the sites record identical channels, at an identical sampling frequency. The monitor will be distributed to participants in the REVAMP arm without in-person set-up instructions, and they will be directed to the REVAMP website to review a video on OSA and PAP treatment developed by the American Sleep Apnea Association and a step-by-step instructional video and brochure explaining HST set-up *(Appendix 2)*. Postal materials will be provided to mail the monitor back to the sleep center using a carrier that allows tracking of the package. Participants randomized to in-person management will be scheduled to meet with a technologist at the sleep center to receive in-person instructions on how to perform the test. During the visit, the subject will be shown the videos and provided with the set-up brochure *(Appendix 2)*. The therapist will explain how to operate the monitor, demonstrate placement of the sensors, and answer the participant's questions. The patient will be given the option of returning the monitor by mail or in-person. Participants in both arms will be provided a 24/7 phone number to contact an on-call PSG technologist at the site's sleep center with any questions about the set-up.

Scoring of sleep testing. The HSTs and PSGs will be scored by an experienced PSG technologist at the respective clinical site and interpreted by a sleep specialist at that site who is blinded to randomization arm and not involved with participant management. The decision not to have centralized scoring of the sleep tests increases the "real world" nature of the project, improves turn-around time, and reduces project cost. On both HST and PSG, events with at least a 30% reduction in respiration for at least 10 sec associated with a 4% or greater oxygen desaturation will be scored as hypopneas. A recent study by the PI that compared scoring of PSGs at 5 sleep centers found an interclass correlation coefficient for AHI of 0.98 using these scoring criteria.[58] To standardize scoring procedures at the sites, a manual that details scoring procedures will be developed. AHI will be calculated as the average number of apneas and hypopneas per hour of sleep on PSG and as the average number of those events per hour of recording on HST.

Long-term follow-up. Following the 3-month follow-up, participants will be scheduled for a follow-up with the sleep therapist every 6 months for the duration of their participation. Participants in the in-person arm will have in-person clinic appointments. Participants in the REVAMP arm will have phone clinic sessions.

Controlling patient wait time. We will monitor patient wait time throughout the protocol to compare the efficiency of REVAMP versus in-person management. REVAMP management will have a fixed schedule as detailed above. In order to match this timeframe, we will set a maximum limit to wait times for participants in the in-person group. Wait time for current in-person care varies widely across sleep centers. In order to enhance the applicability of our results to other sleep centers in the VHA and private sector, we will limit wait times of subjects in the in-person arm so that the initial clinic visit and sleep testing occur within one month of enrollment and the first follow-up clinic visit occurs within the first month of PAP treatment. This represents the current practice of

many private sleep centers. In order not to interfere with the routine clinical care at each site, schedules for research-related phone and in-person clinic sessions will be created. Finding similar PAP adherence and functional outcomes but differences in cost-effectiveness between REVAMP and in-person care when similar wait times are employed makes it likely that even more beneficial effects will occur when in-person wait times are longer. An advantage of the study's design is that we will be able to use our clinical and economic results to construct a decision model to assess the likely effect of delay due to waiting times.

Positive airway pressure treatment. *Participants initiated on PAP treatment will be prescribed a SystemOne RemStar unit (Philips Respironics, Inc.) or an AirSense 10 unit (ResMed, Inc.) set at a pressure range of 4-20 cm H_2O with in-line heated humidifier and wireless technology. The type of unit prescribed will be tracked. A strength of REVAMP is its ability to receive PAP information from both of these major manufacturers and display it on a common platform. The ability to use both manufacturers increases the generalizability of our study. Previous studies, including one by the PI, have validated the accuracy of the AHI determined by the PAP units of both manufacturers, i.e., the measure that will be used to assess treatment efficacy.*[59-63] The use of heated in-line heated humidification promotes adherence.[64, 65] The APAP units will be delivered to the participant by a therapist from a contracted home healthcare company. To create a patient-centered experience, the therapist will custom-fit the participant with the mask interface they prefer and explain how to operate and care for the equipment. All participants will be informed that their APAP unit will record their adherence to treatment and how well the treatment is working to correct their breathing disorder during sleep.

All practitioners will have access to the participant's daily APAP data at the manufacturer's HIPAA compliant, password protected web-site (EncoreAnywhere, Philips-Respironics, Inc *or AirView, ResMed, Inc*). Participants in the REVAMP arm and their practitioners will also be able to view the PAP data on the REVAMP website. When clinically indicated, adjustments to the APAP pressure range will be made either at the time of the in-person clinic visit or by entering this modification on the manufacturer's website that will reset the participant's APAP device by transmitting the information to the home unit via the wireless connection. Participants will be informed if the pressure range is adjusted. Subjects who unable to tolerate APAP or refuse treatment will be seen by a sleep physician for further evaluation and management.

Blinding of research personnel. Given the differences in management of the two groups, the research personnel cannot be blinded to a participant's group assignment. The practitioners delivering care will not be involved with scoring the HSTs and all of the participant questionnaires are self-administered. The practitioners will not have access to the patient satisfaction questionnaire results or to the data collected for Aims 2 and 3.

Formative evaluation (Aim 3). Through mixed quantitative and qualitative methods, we will explore patient- and *practitioner-centered* treatment acceptability and process of the two methods of OSA management. Quantitative: We will evaluate participant acceptance of web-based OSA management versus in-person management by comparing the between-group differences in process indicators, including attrition and therapeutic alliance. *We will evaluate practitioner acceptance of web-based OSA management versus in-person management by comparing the between-group differences in therapeutic alliance.* WAI-SR and CSQ-8 are patient-oriented scales which have been used widely as process indicators in research comparing delivery of treatment via telemedicine versus in-person format.[66, 67] Qualitative: We will conduct semi-structured phone interviews with a subset of up to 45 participants: 15 who complete the first 3-mo web-based intervention, 15 patients who complete the 3-mo in-person management, and 15 patients who drop out of either intervention group. Each patient will be interviewed over the phone by Dr. True who has extensive expertise in qualitative interviewing. The phone interviews will be conducted within one month of the participant completing either the first 3 months of PAP intervention or dropping out of the study. Phone interviews rather than in-person interviews were chosen for a number of reasons: the relatively non-sensitive nature of the interview topic; to decrease burden on patients; and to increase likelihood of successfully completing the targeted number of interviews.[68] A semi-structured interview guide will be developed by Dr. True and revised in consultation with the research team. The interview guide will cover topics including: the participant's experiences with OSA; expectations going into treatment; attitudes, opinions, and experiences with treatment; reasons for dropping out (if relevant); perceived barriers and facilitators to continuing treatment; perceived therapeutic relationship with the provider; and suggestions for making treatment more patient-centered. We anticipate we will reach 'theme saturation' (where no new concepts emerge in the data) after completing approximately 12 interviews with patients from each group.[69] Each interview will be audio-recorded and last between 30-60 minutes.

To explore practitioner-level acceptability of REVAMP delivery of care compared with in-person delivery, as well as perceived barriers and facilitators to implementation of OSA management via REVAMP, we will conduct semi-structured phone interviews with the practitioners who provide OSA care and management for the study. As with the patient interviews described above, each practitioner will be interviewed over the phone by Dr. True using a semi-structured interview guide developed by Dr. True with input from the rest of the research team. The interview guide will cover topics including: the practitioner's background and prior experience with OSA treatment and management; experiences with barriers and facilitators to treating and managing patient's OSA using REVAMP; perceptions about patient's experiences with and attitudes toward receiving care via the REVAMP

platform; and suggestions for improving delivery of OSA care via REVAMP. Practitioners will be interviewed after they have delivered care to a minimum of 20 participants in the REVAMP intervention group to ensure sufficient level of experience with REVAMP and with a variety of participants. Each interview will be audio-recorded and last between 30-60 minutes.

OUTCOME MEASURES

Aim 1 – Outcome Measures (see Appendix 3 for validated questionnaires)

Subject Characteristics Questionnaire is a 15-min self-administered and provides the following information: sleep-related complaints, history of risk factors for sleepiness, past medical history, education level, employment status, current medication use, and an assessment of health literacy.[4]

Functional Outcomes of Sleep Questionnaire (FOSQ-10). The FOSQ-10, a disease-specific questionnaire and our primary outcome measure, will be used to compare functional benefit from APAP treatment.[47] The 10-question instrument has been used in research and clinical practice to measure the impact of daytime sleepiness on activities of daily living. FOSQ-10 has an internal consistency of α = 0.87 and a pre-treatment correlation with the total FOSQ (36 questions) of r = 0.96. After 3 months of CPAP treatment the correlation was r = 0.97. Effect sizes with treatment for both instruments were highly correlated and indicate ability to measure meaningful change.[47] Differences were observed between FOSQ-10 scores for normal controls and OSA patients. The measure is at a 5th grade reading level and takes less than 5 min to complete.

APAP adherence will be objectively monitored by exporting the data from the REVAMP website. Average hours per day of APAP use over the initial 3 months of treatment will be used to test our hypothesis that participants who receive telemedicine-based versus in-person care have similar APAP adherence. In addition to adherence data, the APAP units provide information on treatment efficacy.

Epworth Sleepiness Scale Questionnaire (ESS) assesses perceived daytime sleepiness and will be used to test our hypothesis that participants in the two arms have similar improvements in subjective daytime sleepiness (change in person level disability) on PAP therapy.[48, 49] The respondent rates the likelihood of falling asleep in eight soporific situations using a 4-point Likert scale ranging from "never dozing" to "high chance of dozing". Multiple RCTs report a reduction in ESS with PAP therapy in OSA patients.[5, 23, 70-72]

Medical Outcomes Study Short Form-12 (SF-12) is a 12-item self-administered health-related quality of life questionnaire, with established reliability and validity, providing summary information on physical and mental health status.[51] The physical and mental health scores will be used to compare overall functional benefit between the two interventions. SF-12 uses a Likert scale of 1-3 for physical function; 1-5 for bodily pain, social function, and general health perceptions; 1-6 for vitality and mental health; and a dichotomous scale of yes/no for the presence of role function limitations. The higher the score, the higher the level of health or functioning. SF-12 is widely used to assess health-related quality of life in many populations, including OSA.[5, 70]

Center for Epidemiological Studies Depression Scale (CES-D) is a 20-item standard instrument devised using items from 5 validated depression scales.[52] Likert-style responses for each symptom indicates how often in the past week the subject has experienced the symptom. Scores range from 0 (not at all to 1 day in the week for all 20 symptoms) to 60 (5 to 7 days in the week for all 20 symptoms). The internal consistency of the measure (Cronbach alpha value) ranges from 0.83 to 0.90.

Aim 2 - Measures to be used in cost, utility, and cost-effective analysis (see Appendix 3 for questionnaires)

Medical service use and costs, stratified by whether or not they are related to the diagnosis and treatment of OSA, will be collected for the 6 months prior to enrollment and the entire observation period following enrollment. We will track the cost of all tests, treatments, and equipment (for diagnosis, PAP). Starting in the third quarter of the study (equivalent to the beginning of the second quarter of enrollment) and continuing quarterly for a total of 6 quarters, providers will be asked in the second week of the quarter to report the time they spent managing individual participants during the prior week. Obtaining measurements for a week reduces problems of recall; obtaining them over a year allows capture of changes in physician time that may occur over the study period. The cost of participant travel time will be based on the distance the participant lives from the sleep center. Participants will be asked to recall the time they spent performing activities related to sleep services during their one- and three-month follow-up interviews. We will also collect all costs pertaining to non-OSA sleep-related disorders.

Medical service use and cost data incurred within the VHA will be obtained through access to the Managerial Cost Accounting (MCA) System at the Corporate Data Warehouse (CDW), the VA national data repository. MCA National Data Extracts (NDEs) data originate from various site-level MCA databases and subsystems at the patient level. We will select NDEs to obtain in-patient, outpatient, laboratory, pharmacy, radiology, and surgery services provided to each participant and their associated cost. To gain research access to MCA data, a Data Access Request Tracker (DART) application will be submitted to the National Data Systems (NDS). Once approved for access, the requested information, typically in the format of SQL server output tables, will be provided by the VA Informatics and Computing Infrastructure (VINCI). Fee-based non-VA care on the CDW will be requested from VINCI. Cost and use of Centers for Medicare and Medicaid Services (CMS) by participants

who are CMS beneficiaries will be obtained by submitting a standard request to the VA Information Resource Center (VIReC) to access the VA/CMS repository data that is also available on the CDW. Non-VA and non-CMS services not captured by VA database inquiries will be derived from patient self-reported data collected on questionnaire. The latter service use will be costed out using data that will be derived from the cost data at the respective sites as well as Federal reimbursement schedules when that medical center's cost data are not available. Participants will be asked to provide details about any traffic accidents and we will cost them out using published data on the cost of such accidents. Costs of wireless PAP data transmission, when applicable, will be included in the analysis. Mr. Roberts, our database manager, has strong experience working in VINCI and with SQL server output tables to create SAS datasets.

Preference-weighted QOL. As recommended by the Panel on Cost-effectiveness in Health and Medicine, QALYs will be assessed using preference-weighted generic quality of life instruments.[73] Preference-weighted health-related quality of life will be expressed in terms of the SF-6D (primary preference measure calculated by use of the SF-12 scoring algorithm),[6] and the EQ-5D (secondary preference measure).[8] We propose use of both the SF-6D and EQ-5D for secondary analyses of treatment differences because prior research has indicated that patients with OSA provide significantly different responses to these 2 instruments,[74] While there is little information available for identifying which is a better measure of preference-weighted QOL, either in general or specifically for OSA and its treatment, Schmidlin et al.[74] have recommended use of the SF-6D and the EuroQOL feeling thermometer for the evaluation of OSA treatments.

SF-6D will be used as the primary measure of patient preferences.[6] SF-6D can be derived from the SF-12, which, in combination with data on survival, allows the calculation of QALYs. The SF-6D yields a preference-based single index measure for health from these data using general population values. It focuses on seven of the eight health domains covered by the SF-12: physical functioning, role participation (combined role-physical and role-emotional), social functioning, bodily pain, mental health, and vitality. Scoring of the SF-6D takes into consideration any limitations in the kind of work or other activities as the result of physical health; accomplishing less due to emotional problems; bodily pain and its interference with normal work; nervousness, depression, and energy level; and interference with social activities due to physical or emotional problems. Research has shown that a meaningful health state classification measure, the SF-6D, can be created by applying a scoring method developed by Brazier and colleagues.[75, 76]

EQ-5D will be used as a secondary measure of patient preferences.[7] The self-report questionnaire includes the respondent's self-rated health status on a vertical graduated (1-100) visual analogue scale. In addition, a multi-attribute utility scale comprises 5 dimensions of health (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) and produces a single index score for each state of health.[7] Each dimension comprises 3 levels (no problems, some/moderate problems, extreme problems). A unique score is calculated for the EQ-5D, by applying a scoring method developed by Shaw and colleagues.[8]

Aim 3 – Formative evaluation (see Appendix 3 for questionnaires)

Working Alliance Inventory-Short Revised (WAI-SR) is the 12-item short form version of this widely used instrument assesses three key aspects of the therapeutic alliance: (a) agreement on the tasks of therapy, (b) agreement on the goals of therapy and (c) development of confidence, trust, comfort, and acceptance.[9] WAI-SR has good internal consistency (alpha ranges from 0.89-0.92) and convergent validity.

Client Satisfaction Questionnaire (CSQ-8) is an 8-item, self-administered measure of client satisfaction with services.[10] The CSQ-8 is unidimensional, yielding a homogeneous estimate of general satisfaction with services. CSQ-8 has been extensively studied, and while it is not a measure of a client's perceptions of gain from treatment or outcome, it elicits the client's perspective on the value of services received.

STATISTICAL ANALYSIS

Sample size, and power. Our primary aim is to demonstrate that the efficacy of REVAMP management is not clinically inferior to in-person care. Efficacy will be measured as the change in FOSQ-10 score from pretreatment to 3 months following initiation of APAP treatment (Hypothesis 1a). In the recent study by the PI of patients with suspected OSA (N=296) randomized to either home testing or standard in-laboratory testing, mean (SD) functional outcome score improved 1.74 (2.81) in the home group and 1.85 (2.46) in the in-lab group (both $p<.0001$)[5] , and improvements in FOSQ in excess of 3 have been reported.[24] Increases less than 1 point are not clinically significant.[77] Assuming that an increase of more than 1 point in mean FOSQ score would indicate an important clinical difference, then the non-inferiority hypotheses for FOSQ score may be expressed: H_0: Δ_R-$\Delta_P \leq$ -1.0 vs. H_a: Δ_R-Δ_P > -1.0 where ''R" represents REVAMP management and "P" is in-person management. Assuming a common SD=2.80, 114 patients/group are required to achieve at least 85% power to reject the null hypothesis, i.e., find that the change in FOSQ score among patients receiving the telehealth-based management is > 1 point lower than that in patients assessed in-person.

Using a similar approach to estimate the sample size needed to test non-inferiority of APAP adherence (Hypothesis 1b), mean (SD) PAP use of participants enrolled in our REVAMP pilot study was 5.3 (2.2) hours/day (see Preliminary Results). Based on a non-inferiority delta of -0.75 hr/day and assuming a SD=2.2, the 114

patients/group needed in Aim 1 will achieve an 80% power to reject the Aim 2 null hypothesis, i.e., PAP adherence among patients receiving REVAMP management is > 0.75 hr/day lower than that in patients assessed in person. We note that the projected sample size for Aim1 is the same as that achieved in our previous study which demonstrated clinical non-inferiority of home vs in-lab testing using the same outcomes.

Based on current referrals to the *3 clinical sleep centers*, we anticipate that 20% of subjects enrolled will not have OSA. Adding to this estimate a 20% drop-out rate of patients diagnosed with OSA and initiated on APAP, we need to enroll a total of 356 subjects, i.e., approximately119 subjects per clinical site. We are confident that we will be able to achieve this goal. Given our clinic and personnel resources, we plan to enroll 6 patients per month per site. We anticipate that it will take about 2 years to enroll the participants (see Gant chart). *Thus, we will have at least 9 months of post treatment cost data on all participants for our cost-effectiveness analysis, while the longest will be about 2 years. The longer follow-up should provide the opportunity to observe potential long-term cost benefit and outcome differences between the testing methods.*

Descriptive comparisons at baseline. Descriptive statistics will be computed for all outcomes; distribution checks will be performed using histograms, quantile plots, and goodness-of-fit tests. Sample characteristics will be summarized using mean (SD) for continuous outcomes and frequency and percent for categorical variables. Descriptive analyses will be performed to characterize the treatment groups separately by clinical center and to confirm that the randomization resulted in no clinically significant treatment group differences at baseline. Partial correlations (controlling for site and group) between baseline factors and outcome measures will be computed to identify potentially important covariates. The primary non-inferiority tests will be repeated, controlling for baseline factors that are found to have both clinically significant treatment group differences and evidence of association with either 1-month or 3-month change in FOSQ score.

Statistical methods for non-inferiority hypotheses. Aim 1 will determine if REVAMP management is not clinically inferior to in-person treatment, in terms of the difference between baseline and post-treatment FOSQ score (Hypothesis 1A), and mean APAP use (Hypothesis 1B). *To test the Aim 1 hypotheses, we will use a modified intent to treat analysis set, i.e., participants with a diagnosis of OSA who were treated with PAP with or without supplemental oxygen. This strategy will avoid potential bias in the non-inferiority analysis of treatment effect.* To demonstrate non-inferiority for Hypothesis 1a we will test the primary non-inferiority hypothesis: H_0: Δ_R-$\Delta_P \leq -1$ vs. H_a: Δ_R-$\Delta_P > -1$; where Δ_R and Δ_P are the mean FOSQ-10 change scores from baseline to 3 months in the REVAMP (R) and in-person (P) groups, respectively, and where a positive change score indicates improvement. In these equations we have explicitly included the value of the pre-specified non-inferiority margin, δ=1, for change in FOSQ score. The hypothesis will be tested by using analyses of covariance to construct least squares estimates of the (adjusted) group differences in mean values along with their 90% confidence intervals (CI). The lower bound of the 90% CI is equivalent to the lower bound of a 95% one-sided non-inferiority CI. If the lower bound exceeds the *a priori* non-inferiority delta, then the null hypotheses of inferiority can be rejected at α=0.05.

The primary analysis will use mixed effects models for longitudinal data which make use of all available scores for each participant at each measurement time and permit adjustment for inter-group differences that are not balanced through randomization. All analyses of mixed models will account for clinical site. For the change in FOSQ score, baseline FOSQ score will also be included as an *a priori* covariate. Factors that are pre-specified as being possible confounders in the association of treatment assignment and outcome are age, sex and race. We anticipate little confounding from these patient-level factors given the blocked randomization scheme to ensure balance between treatment groups. The site by treatment group interaction will not be included and this decision reflects recommendations that differences among treatment group means should be estimated weighting sites according to the amount of information they provide. With this weighted estimate, each within-center estimate is given a weight based upon its precision.[78] A series of mixed models is proposed; each outcome identified in the aims will be rotated individually in the modeling process. We assume that change outcomes will be normally distributed but will check the distribution of each outcome before modeling, and, if necessary, use a generalized linear mixed model to accommodate alternative distributions.

A similar analysis will be conducted to test Hypothesis 1b: Ho: μR-μP $\leq$ -0.75 hr/night vs. Ha: μR-μP $>$ -0.75 hr/night; where μR and μP are the mean APAP hr/night of patients receiving REVAMP (R) versus in-person (P) management. If the lower bound of the 95% one-sided confidence interval exceeds the non-inferiority delta of -0.75 hr/night, then the null hypotheses of inferiority can be rejected at α=0.05.

Multiple comparisons. No adjustment for multiple comparisons will be performed to account for testing of two primary hypotheses, since rejection of both non-inferiority hypotheses are necessary to conclude that at-home assessment is not clinically inferior relative to in-lab assessment.

Analysis of secondary clinical outcomes in Aim 1. Secondary outcomes in Aim 1 include changes from baseline to Month 3 in scores of the ESS, SF-12 (physical and mental health components), and CES-D questionnaires. For each secondary outcome, analogous analyses of covariance models will be used to determine least squares adjusted treatment group differences in mean change along with 95% confidence

intervals since non-inferiority deltas are not pre-specified. Baseline values of the outcome measures will be included in their respective ANCOVA models. The above results will be interpreted as hypothesis generating.

Missing data methods. The mixed model approach uses information in the covariance matrix of predictors and outcomes over time to implicitly impute missing 3-month FOSQ values. Of note, participants will complete the assessment battery at 1-month following initiation of treatment. Therefore, subjects missing month 3 results will not be excluded. These imputations are valid if the likelihood of missingness is independent of treatment group, conditional on observed data included in the model, i.e., if missing data are missing at random (MAR). For comparison with earlier studies, we will also perform the analyses using the last observation carried forward method.[5] Finally, in sensitivity analyses, we will evaluate the impact of departure from the assumed MAR mechanism by using multiple imputation (MI) to implement a selection model approach.[79] In this way, we will be able to able to compare estimated effects based on different assumptions about the missingness mechanism, and to assess the sensitivity of our estimates to these assumptions. If data on predictors is incomplete, we will use multiple imputation methods (as implemented in SAS PROC MI and MIANALYZE) to generate multiple imputed complete datasets, perform analyses, and then combine analysis results to produce estimates that are asymptotically unbiased, with standard errors that account for imputation uncertainty.[80]

Modified intent-to-treat and per-protocol populations. Primary analysis for non-inferiority will be based on all randomized participants initiated on APAP with at least one FOSQ follow-up score. In a secondary analysis, the non-inferiority test will be repeated using a per protocol cohort (patients who have been initiated on APAP and have non-missing endpoints at month 3), since protocol violations can add noise making it more difficult to distinguish between the two active treatment groups.

Statistical methods for Aim 2. *Aim 2 will determine if REVAMP management of patients with OSA is cost-effective compared with in-person care. The perspective of the analysis will be that of the VA. While the VA has the general well-being of veterans as its goal, we will focus our attention on health care costs, in part because VHA patients with OSA are older and less likely to work. The analyses for Aim 2 will be based on all randomized participants, as this is the target population for the evaluation of Aim 2 and Aim 3 outcomes related to management of this patient population. We will perform a secondary analysis using the modified intention to treat sample (Aim 1 analysis set). The primary outcome measure in Aim 2 will be the difference in sleep-related cost between the two delivery modalities. Our hypothesis is 2a) Ho: $\Delta C = 0$ vs HA: $\Delta C<0$. Based on data from our prior study of home vs in-lab diagnosis,[81] we assume average 3-month sleep-related costs of $1755 (SD, 745) for the home group and of $2538 (SD, 1065) for the in-lab group. Resulting power is 1.0 for sample size of 114 per group and remains greater than 0.9 if, when including participants without OSA on HST and participants with CSA/CSR, the means are reduced by 25% and the SDs are increased by 25%.*

Secondary outcome measures include the difference in total costs, the difference in QALYs, and the ratios of the differences in cost and QALYs. The secondary hypothesis for the cost-effectiveness ratio is 2b) Ho: $\Delta C/\Delta Q > 100{,}000$ vs Ha: $\Delta C/\Delta Q \leq 100{,}000$. As with the efficacy and non-inferiority analyses, the cost difference (ΔC) and the preference score difference (ΔQ) will be estimated using mixed effect models for longitudinal data. QALYs will be constructed by taking the area under the quality-adjusted survival curve.[82] The cost-effectiveness ratio ($\Delta C/\Delta Q$) will be calculated by dividing the estimated difference in cost by the estimated difference in QALYs. Confidence intervals for the cost-effectiveness ratios will be calculated using Fieller's theorem method.[83] A point estimate less than 100,000/QALY and a 90% confidence interval that excludes $100,000 will be evidence that we can reject the null hypothesis that the incremental cost-effectiveness ratio is not less than $100,000.

Costs, QALYs, and cost-effectiveness ratios will be calculated for costs and effects during the first 3 months following initiation of APAP treatment and for the full period of follow-up (maximum, 2 years). The study design (follow-up extending from enrollment until 3 months after last patient is enrolled) introduces substantial administrative censoring that – as with the primary analysis -- will be addressed by use of a mixed model approach (i.e., patients will have different lengths of follow-up, and thus different costs and QALYs, simply because they were enrolled at different times). Because this censoring is due primarily to administrative reasons, it is generally considered to yield data that are missing completely at random (MCAR). While use of the electronic medical records for the collection of cost data will lead to limited missing cost data, there may be censoring of cost data for services outside the VHA and of the QALY measures due to loss to follow-up. If such drop-out is related to disease severity or to outcomes after randomization, it generally would be considered to yield data that are either missing at random (MAR) or that are non-ignorably missing.

Statistical methods for Aim 3, Quantitative component. Aim 3 will compare patient-centered treatment acceptability and process measures of the two management approaches. We will evaluate patient acceptance of telemedicine-based care by comparing the between-group differences in attrition, therapeutic alliance, and satisfaction with care. The primary quantitative objective outcome measure in Aim 3 will be percent of patients who drop out of each treatment. The primary subjective outcome measures will be change in the WAI-SR and CSQ-8 scores. For this exploratory aim we will estimate mean values and confidence intervals. Mixed effects models will be used to estimate the between-groups difference in means and confidence intervals, for each of

the outcome measures. Results will be ascertained via ITT and per-protocol procedures as previously discussed. A series of mixed models is proposed for each outcome identified in Aim 3 (WAI-SR, CSQ-8).

Qualitative data analysis. Interviews will be digitally recorded with the participant's permission, transcribed by a professional transcriptionist, and prepared as primary documents to be loaded into Atlas.ti software for data management and analysis. Atlas-ti[84] is a software program designed to assist in the management and analysis of qualitative data, including primary documents (e.g., transcripts), secondary notes (e.g., memos written by the investigators concerning methodological issues), and analytical notes (e.g., emerging theories to be tested through further data collection and analysis). Atlas is particularly suited to the constant comparative method most closely associated with grounded theory, adapted for use in healthcare and health services research.[85] Analysis of patient interviews will be guided by Aim 3: to identify patient perspectives, attitudes, and preferences regarding REVAMP versus in-person management of patients with OSA, as well as barriers and facilitators to participation in either clinical pathway. *Similarly, analysis of provider interviews will focus on identifying providers' attitudes and perspectives on acceptability of delivering care via REVAMP vs. in-person, and perceived barriers and facilitators to implementation, spread, and maintenance of REVAMP as a means to deliver treatment for patients with OSA.*

Analysis will follow a series of systematic steps in order to provide a structured approach to ground the flexible, exploratory nature of qualitative research and ensure transparency of the process for examination by other researchers in the field.[86] An initial set of codes for each type of interview data (e.g., patient, provider) will be developed and defined after reviewing the first few transcripts for each type of interview (open coding). The codes will be revised and refined through discussion involving research team members. Any disagreements in coding will be resolved through discussion with the research team and once the final coding scheme is standardized, all transcripts will be coded by a research assistant overseen by Dr. True. Examination of coded text across transcripts and relationships between codes (axial coding) will identify recurrent themes and categories related to Aim 3. Separate codebooks will be developed for coding of each type of interview data *(patient, provider),* but analysis and summary of data will consider themes that emerge within each group (e.g. patient-specific themes, provider-specific themes), as well as themes that emerge across both groups.